DOCUMENT RESUME

ED 119 498 FL 007 491

AUTHOR Matluck, Joseph H.; Mace-Matluck, Betty
TITLE The Multilingual Test Development Project: Oral Language Assessment in a Multicultural Community.
PUB DATE Mar 75
NOTE 13p.; Paper presented at the National Conference of the Teachers of English to Speakers of Other Languages (Los Angeles, California, March 1975)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage
DESCRIPTORS Bilingual Education; *Bilingual Students; Cantonese; *Child Language; Communication Skills; Cultural Differences; Diagnostic Tests; Elementary Education; English (Second Language); Fundamental Concepts; Language Instruction; *Language Skills; *Language Tests; Mandarin Chinese; Oral Communication; Phonology; Second Language Learning; Spanish; Tagalog; *Test Construction; Testing; Vocabulary
IDENTIFIERS Ilokano

ABSTRACT

This paper discusses a series of oral proficiency tests in six languages developed under the auspices of the Center for Applied Linguistics and the Seattle Public Schools District. The prototype is an English test, totally oral, for use in grades K-4, designed to: (a) determine the child's ability to (1) understand and produce the distinctive characteristics of spoken English, (2) express known cognitive concepts, and (3) handle learning tasks in English; and (b) provide placement and instructional recommendations for alternate programs such as special English instruction and bilingual education. Comparable tests with similar objectives were then developed in Cantonese, Mandarin, Tagalog, Ilokano, and Spanish, reflecting some of the largest of the forty non-native-English-speaking groups in Seattle schools. Development of the tests was based on identifying basic learning concepts that children must handle in order to perform in a school setting. These concepts include identifying, classifying, quantifying, interrogating, negating and showing spatial, case and temporal relationships. The grammatical manifestations of a language that a child must handle to perceive or to communicate these concepts were then determined. The construction of the tests and the scoring methods are discussed, as well as cultural and linguistic differences encountered in the preparation of the tests. (Author/CLK)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED119498

THE MULTILINGUAL TEST DEVELOPMENT PROJECT: ORAL LANGUAGE ASSESSMENT IN A MULTICULTURAL COMMUNITY

Joseph H. Matluck
The University of Texas
at Austin

Betty Mace-Matluck
Seattle Public Schools
Seattle, Washington

FL007491

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

If we are to design adequate instructional programs in the schools of any multilingual community for children whose native language is not the dominant language of the country, we must know the level of the child's audio-lingual proficiency in both languages. To do this, we must have tests in each of these languages which accurately measure that proficiency.

Several years ago the Seattle Public Schools became deeply concerned with this problem. Because of a series of factors involving immigration patterns and movement of populations, they were faced with a great number of children at all levels of instruction who spoke a language other than English and whose control of English was so inadequate that they could not function in the regular school program.

This has also been the special concern of the Center for Applied Linguistics in Washington, D.C., heavily involved in applied linguistics, in TESOL, and more recently in bilingual education.

These concerns were brought into sharp mutual focus a few years ago by a research project, conducted in Seattle (Mace, 1972), which provided a sociolinguistic profile of these children in the Seattle Schools. This research indicated in the first place, that there were forty languages represented among the students registered in Seattle's elementary schools, and in the second place, the sad state of the art in the field of oral proficiency test construction.

Unfortunately at this time there are still few good tests available, and of these, few avoid the many pitfalls which are so evident in existing tests. Few are designed for young children; precious few are completely

oral-aural; very few really test much of anything; few exist in any language other than English; and even fewer make an adequate distinction between receptive and productive skills, between speech and writing, between phonology and grammar, etc. Most of them use only one mode of testing and many make no attempt to differentiate an urban setting from a rural one nor to examine the content and the visuals carefully enough to free them from cultural bias.

The Multilingual Test Development Project, on which we have been working for the past year in Seattle, under the auspices of the above-mentioned organizations and with the generous assistance of the Center for Research Services for Multicultural Education, was an attempt to correct some of these ills. The MAT-SEA-CAL Oral Proficiency Tests (Matluck-Matluck, 1974) in English and five other languages were developed for use with young children. The prototype is an English test, totally oral, for use in the early grades (K - 4), designed to determine the child's ability to understand and produce the distinctive characteristics of spoken English, to express known cognitive concepts and handle learning tasks in English, and to provide placement and instructional recommendations with respect to alternative instructional programs such as bilingual education, special English instruction, etc.

With the help of native informants, we then developed comparable tests with similar objectives in Cantonese, Mandarin, Tagalog, Ilokano and Spanish, languages which reflect the five largest of those forty non-native-English-speaking groups in attendance in the Seattle Public Schools.

Modes. Each test uses various modes of assessment: Listening Comprehension, Sentence Repetition, Structured Response, and incorporates a wide range of language features which are broken down for analysis and scoring into 1) phonology and 2) grammatical and lexical structures.

Communication Concepts. The construction of tests in only one language

can be made relatively easy by the identification and selection of its most characteristic grammatical features. If we add one other language and attempt to develop two tests - not parallel but comparable - the task gets much more difficult because we must change the order of presentation in order to accomodate the grammatical structures of both languages. The moment we add a third language, or a fourth or fifth or sixth, and try to use the same type of construction, we are confronted with complete chaos. Obviously some other format had to be developed in order to construct, simultaneously, comparable tests in six languages.

We approached the development of the tests from the point of view of identifying some basic communication concepts that a child must handle in order to perform in a school setting, i.e. the skills of identifying, classifying, quantifying, interrogating, and negating and of showing important relationships such as spatial, case, and temporal. We then determined the grammatical manifestations of the language - English for example - the child must handle in order to perceive or to communicate these concepts. The other languages express the same concepts as does English, but each one almost always does it in different ways, i.e. with different language manifestations.

A couple of examples:

1) In order to quantify, English inflects nouns (/-s/, /-z/, /-ɨz/, but not adjectives nor verbs (except very partially in one tense only). Spanish inflects nouns, adjectives, and verbs for plural. Mandarin and Cantonese inflect none of these for plural, but use other means to pluralize nouns: numerals and certain adjectivals; Tagalog and Ilokano inflect neither nouns nor verbs for plural but do inflect adjectives and use pronouns and certain adverbs to indicate the plural subject of verbs.

2) In order to show temporality, English does it almost entirely through the verb structure, as does Spanish, but in different ways (i.e., the Spanish

verb inflectional system is much more extensive). In Mandarin and Cantonese verbs don't reflect temporality at all, and this concept is signalled by adverbs of time and some particles which change the status of a condition, occurance, or action. Tagalog and Ilokano do it partially through the verb structure (but in a different way from English and Spanish), and partially through adverbs of time.

Master charts. Each section of the test is accompanied by a master chart whose coordinates reflect the relationship between the concept and the language manifestation for each grammatical item tested.

Vocabulary inventory. The English test consists of 222 different words, all of which appear on high-frequency word lists of English. In addition, the test requires the student to generate a minimum of about forty additional words, all but two or three of which are also included in the high-frequency word lists. The vocabulary inventories of the other languages have similar range and proportion.

Administering of tests. There are eighty-one sentences in each test. The estimated time for administering the entire test is 25 - 40 minutes (depending on the language being tested), and no one section takes longer than 15 minutes. Part I is designed for individual or group administration; Parts II and III for administration to one student at a time. Each section can be used independently. Part III, especially, can be used as a fairly good quick overview of the structures the child commands, including the phonological.

Phonology. The phonological items tested in each of the languages provide a fairly complete phonological inventory of that language, and we have worked out a scoring system that would neither be too cumbersome for the average teacher to handle, nor would it weigh pronunciation too heavily in the total score. There are two independent percentage scores, one for phonology and one for "structure".

Tapes. Two of the three parts of the test in each language have a taped stimulus; the stimulus for the third (Structured Response) is supplied by the examiner. Of course, native speakers of the language being tested are used to tape, administer and score the test in each language. All of the student's oral responses are taped and all items in the test have a visual reference accompanying and reinforcing each aural stimulus, with the exception of a small "Response-to-Commands" section.

Examiner's Handbook. The heart of the system is, of course, the Examiner's Handbook which contains, among other things, complete instructions for administering and scoring the tests. For each section of the test there is a sentence sheet, sample visuals, a sample tape script, a vocabulary inventory, a phonology chart, and to facilitate scoring, a set of recap sheets. For each sentence of the test, there is a recap sheet containing: 1) the stimulus sentence; 2) the concept and the language manifestation involved in each item being tested in that sentence; 3) each test item extracted from the sentence and placed in the appropriate box (phonology, morphology, syntax, or vocabulary); 4) a check-box identifying those sentences which exemplify "classroom style"; and 5) at the bottom of the page, a reduced version of the visual used for that sentence.

Scoring and diagnostic analysis are accomplished by the use of prepared charts which require the scorer to simply circle the number corresponding to the items missed. Scoring is objectivized by a system which ignores all but the particular portion of the sentence being tested. Percentage scores of each section of the test are averaged for the final composite percentage score and a line graph is supplied which allows the teacher to obtain an overall visual diagnostic profile by simply circling the number in the graph corresponding to the child's score on each concept.

Diagnostic profile. The scoring procedure we designed supplies, for

diagnostic and planning purposes, a profile of 1) the student's ability to handle each concept, and 2) his areas of strength and weakness. Since all of the tests are comparable in their organization and structure and in linguistic content, by administering the test in English and in the child's native language, we can compare the profiles and determine whether the child's problems are linguistic or perhaps conceptual in nature.

A comparison of the superimposed profiles can also indicate the child's language dominance in the areas tested, and can provide useful information as to placement, and curriculum and program planning for that particular child.

We have just begun extensive field testing both in Seattle and in several other areas, and we hope to test at least 500 children in two languages prior to the publication of a revised edition early this summer.

Language-and-culture. Working on these tests and with native informants of all these languages has been a very rich and rewarding experience for us, both culturally as well as linguistically. Most sociolinguists agree that language is the best representation of a culture. But again, too often the myth of lexical equivalence is applied and the uninitiate believe that language differences are manifested by the simple substitution of one word for another, or else by that wonderful inter-language invention of desperate writers of foreign-language grammars called "idioms", and that all we have to do is learn these words and idioms and !Voila! We communicate! And, in reverse, that all any foreigner has to do to communicate with us in our language is to do the same - just substitute our words and idioms for his.

Well, in the first place, the same words are often just not there, and the so-called "idioms" are not idioms at all, but radically different ways of organizing thoughts and ideas, that is, different grammatical structures. There are very intelligent societies with very sophisticated languages who

start their sentences, not with the subject, but with the predicate; others who don't pluralize their nouns; others who don't conjugate their verbs; still others who don't have or don't use any copulative verb (such as our to be), and so on ad infinitum. When people who speak such languages try to learn English and have trouble, it's not because they're unintelligent or primitive or lazy or unmotivated, but rather they're confused -- just as you'd be if you had to learn their language and their culture.

I'd like to discuss some of these with you, along with just a few of the many manifestations of cultural differences which we have run across in our research in connection with the preparation of the tests. If we start by looking at linguistic categories, we find that:

a. Many languages, including most Asian groups, pluralize in a much less primitive way than we do with forms for more than one being divided into duality and plurality categories.

b. As a rule, non-Indo-European languages do not inflect nouns for plural, nor do they usually inflect verbs for tense-person. The noun in Chinese has no article, no inflection and is classified by what for us is an unfathomable system - by size, shape, use, quality, texture, etc., and these are embodied in a measure word which precedes the noun.

c. In the pronoun classification area, English has sex differentiators; Spanish has grammatical gender differentiators; Chinese has neither, but can classify by a measure-word system; Filipino does not differentiate sex, but does make animate/inanimate distinctions.

d. In the English tag system, surely the craziest ever invented by man, the tag is affirmative if the statement is negative and vice versa. Most languages handle the corroborative interrogative with a simple "true?" or the equivalent, and Asians solve our system in a way which is very logical for them but most disconcerting for us, by simply ignoring the negative or

affirmative reversals in the tag and agreeing or disagreeing with the statement. Thus, "Yes, I did." or "Yes, I didn't."

e. Asian languages normally eliminate the copulative in favor of some type of stative verb system: thus, "Pretty girl Lily", "Salesman Tony". Many native-English speakers are tempted to associate this type of error in the speech of a non-native-English speaker with the linguistic judgement of "pidgeon, primitive, ignorant, baby talk" -- sometimes indeed, carrying this to the ridiculous extreme of judging <u>him</u> and <u>his</u> language as ignorant, primitive, etc.

f. Verb structures present all kinds of complicated differences. In Filipino, normal sentence construction has the predicate precede the subject (or topic) in a manner which seems to us very similar to our passive voice. There is a strong non-correspondence of time, tense and aspect from language to language. Spanish has a highly developed system for time; Chinese never inflects verbs for time but only for change of status and uses adverbs to signal time differences; and Chinese uses verbs to express spatial relationships; Filipino does this to a lesser degree (locative focus requires a special verb class); and Filipino verbs have time distinctions only between action begun and not begun. Filipino languages have a tremendously complicated system of affixes for their many, many verb classes and an equally complicated and varied system of focusing; that is, what the speaker wants to be highlighted becomes the topic of the verb, the difficulty for us being that each time the focus is changed, a completely different class (or conjugation, if you will) of the verb is required, with completely different forms (prefixes, suffixes and infixes), and the focuses vary from actor to object to locative to benefactive to causative to instrumental for a total of at least nine different morphological classes and many sub-classes.

g. The Spanish gender system, unique among these six languages, repre-

sents a gigantic learning problem for the other five groups; and for the Spanish and Filipino speakers especially, our system of inflecting the first of two nouns to show possession (instead of their "the ear of the dog" pattern) represents a three-pronged learning challenge (syntactical, morphological, and phonological) which is equally gigantic: word order, inflection and consonant clusters. And the English prepositional system is a mystery to Asians who really don't have prepositional systems at all, but who use directional and locational markers or directional and locational verbs.

We can also look at each culture's view of reality from the starting point of the reality itself -- as manifested in our tests by our visual materials and their production -- and which often necessitated compromises of varying degrees and in varying directions.

Some cases in point:

a. In order for a Mandarin to recognize a building as a school building, it needs a fence around it; for the Cantonese, it must be a one-story building; the Filipinos need a wall; all seemed to need a flagpole. For the Chinese to recognize children as schoolchildren, they need to have, not books in their hands, but a knapsack across their shoulders. Our policeman first turned up wearing short pants and with his holster on his left side (our artist was a young Cantonese woman from Hong Kong).

b. Dolls can present all kinds of problems across cultures. We used a little doll's bed in one visual - with some trepidation, but we used it. Our fear was the total lack of familiarity with such an entity which might exist on the part of a child belonging to a culture in which nobody sleeps on a bed. And in Mandarin the word for doll is "foreign baby" and the dolls are blonde and look very western (they now have native-looking dolls, but the name stuck). In Filipino dolls are not babies, but adult figures, since

children are treated as adults, invited to adult affairs, learn adult dances (there are no children's dances as such). And the child is not called a child but a young man or a young woman (there is no separate word for "boy" or "girl").

c. A cat with a bandage on its ear and playing with a ball are difficult concepts for some groups because many cultures simply don't get toys for cats and dogs and because these pets are just not that important to people. Thus, it may be very difficult for a Chinese to have to say the equivalent of "What happened? Did he have a fight?" with respect to a cat. And in Hispanic cultures, until very recently, there were no S.P.C.A.'s because animals are simply on a different level of importance.

d. A little boy at an easel is not a familiar experience for a Chinese child. Our paints and brushes are not familiar to him. Therefore there is no connection of cause and effect between the child crying, brush in hand in front of the easel, and the spilled can of paint on the floor (the Chinese child doesn't use that kind of paint). An inference that might be drawn here which could be of great importance to the whole field of testing is that the Chinese child, if asked to make that kind of cause-and-effect relationship, would probably miss that item on our I.Q. tests.

e. It might be very difficult for native-speakers of Chinese to say the equivalent of "He got a bike for his birthday", since something as expensive as a bicycle would never be given as a birthday gift -- perhaps as another kind of present or as some kind of reward, but the birthday "gift" is traditionally a little red envelope with a token coin inside. In Catholic countries "Christmas presents" might be a difficult item, since presents come, not at Christmas but at the Epiphany.

f. In Mandarin it is at least unusual, if not inconceiveable, for a

mother to be cutting her daughter's hair; not so in Cantonese. And Filipino has a different verb for cutting hair than for ordinary cutting.

g. Food presents many problems. Sandwiches and bread are unknown to many cultures; in some rice is the base food. Meat is used very sparingly by some. And what are consummate delicacies for some are disgusting and nauseating to others.

The examples we have touched upon represent only some of the many sociolinguistic problems inherent in multilingual testing and the concomittant dangers of cultural bias in either the visuals or the language employed, and they illustrate, we think, how closely related and inseparable are a language and its culture.

REFERENCES

Mace, Betty J. 1972. "A linguistic profile of children entering Seattle Public Schools kindergartens in September, 1971, and implications for instruction." Master's Thesis, Austin, University of Texas at Austin.

Matluck, Joseph H. and Betty Mace-Matluck. 1974. MAT-SEA-CAL oral proficiency tests: English, Spanish, Cantonese, Mandarin, Tagalog, Ilokano. (Field-test edition in six volumes), Seattle Public Schools, Seattle, WA, Center for Applied Linguistics, Arlington, VA.